Tool Box and Resources

- **Cyberinfrastructure "books"**
    - [Fourth Paradigm: Data-Intensive Scientific Descovery](#) (text book).
    - [Frontiers in Massive Data Analysis](#) (text book) Note: You can get [pdf from NIST](#) as well

- **Introduction to Linux**
    - [Cornell Virtual Workshop (good for novice)](#)
    - [Unix and PERL for biologists](#) (good for novice)
    - [Unix Quick guide](#) (good for novice)
    - Despite its (scary) name, this is an excellent no-nonsense tutorial! http://cli.learncodethehardway.org/book/
    - GNU [Core utils](#) for working with files
    - [LinuxCommand](#) website is good place to get started on some command-line-fu.

- **Programming:**
    - Really good thoughts on programming language environments, learning, and thinking
        - http://worrydream.com/LearnableProgramming/
    - Text editing in Linux:
        - Pico: For the beginner. Pretty simple and gets the job done
            - http://www.cs.colostate.edu/helpdocs/pico.html
        - VI: your best, complicated friend for almost quickly editing text files:
            - The command 'vimtutor' that can be found on popular Linux Distros is a quick way to get acquainted with vi(m).
            - http://heather.cs.ucdavis.edu/~matloff/UnixAndC/Editors/ViIntro.html
            - VIM Cheatsheet: https://cdn.shopify.com/s/files/1/0165/4168/files/preview.png
            - VIM Video Tutorials: http://derekwyatt.org/vim/vim-tutorial-videos/
        - Eclipse: When you want a full blown IDE
            - http://www.cs.colostate.edu/helpdocs/pico.html
        - PyCharm: Some folks really like this Python IDE
            - https://www.jetbrains.com/pycharm/
    - Programming for kids: (really neat way to start)
        - http://scratch.mit.edu/

- **Linux Foo:** http://www.linuxquestions.org/questions/linux-newbie-8/what-is-foo-468303/
    - tmux: (awesome virtual console manager)
        - Let's you keep a "screen" active while you are logged out of the system
        - http://tmux.sourceforge.net/
    - screen: (level 1 virtual console manager)
        - http://www.gnu.org/software/screen/

- **Code Versioning:**
    - GIT concepts simplified (from the ACM club's email): It covers a lot of git concepts and explains them all graphically.
        - http://gitolite.com/gcs/index.html

- **Cloud Data Storage:**
    - iPlant Data Store: https://pods.iplantcollaborative.org/wiki/display/DS/Data+Store+Quick+Start

- **Virtual and Cloud Computing:**
    - [Virtual Box](#): Virtual machines on your machine.
        - [Documentation](#)
    - Amazon: EC2 (requires $$$)
        - http://aws.amazon.com/ec2/

- **Linux System Administration (yes you need it for the cloud):**
    - http://library.linode.com/using-linux/administration-basics
    - http://www.linuxtraining.co.uk/download/new_linux_course_modules.pdf

- **HPC:**
    - Training/learning material for software resources on XSEDE (including paraview, linux intro etc)
        - https://portal.xsede.org/online-training
    - [UA HPC getting started page](#)
        - Batch Queue informationhttp://rc2.webhost.uits.arizona.edu/hpc-htc/using-systems/batch-system
    - Future Grid:
        - Getting an account instructions are at:https://portal.futuregrid.org/manual/gettingstarted
        - Note: FutureGrid is no longer a publicly funded project, but they have made some of their resources available for this course.
- **Example Big Data:**
    - [Backblaze hard-drive failures](#): https://docs.backblaze.com/public/hard-drive-data/2014_data.zip (560.0 MB ZIP file, 3.02 GB on disk, 365 files.)